

TITLE OF THE INVENTION

Mostly Concurrent Garbage Collection

FIELD OF THE INVENTION

5 The present invention relates to garbage collection in computing environments in general, and in particular to concurrent mark-and-sweep garbage collection techniques.

BACKGROUND OF THE INVENTION

10 Garbage collection techniques for use in computing environments are well known. In mark-and-sweep garbage collection, an object connectivity graph is traced from root objects to all their referents and their descendents (i.e., all objects “reachable” from root objects) in an object population, with all traced objects assumed to be in use and marked as such. Unmarked objects are assumed to be not in use and are made
15 available for reallocation. In concurrent mark-and-sweep garbage collection, object tracing and marking is performed in parallel with the operation of an executing program or process, referred to as a mutator. Since the object connectivity graph may be modified by the mutator while the collector traverses from object to object, a write barrier is generally employed by the mutator to record locations of object references that
20 it modifies. Typically, objects are grouped into cards which the mutator marks if it modifies an object reference on a card. Objects on marked cards are then retraced by the collector during a process referred to as card cleaning in order to ensure that all reachable objects are marked.

 While object tracing and card cleaning may be executed in parallel with the
25 operation of the mutator, a card may be marked at any time, including while it is being cleaned or immediately thereafter. Therefore, to ensure correctness, all computing processes are periodically suspended with the exception of the garbage collector which then cleans all marked cards and traces the object connectivity graph from unmarked root objects, ensuring that all reachable objects are marked. However, the “stop-the-
30 world” suspension of all processing is often disruptive to processing flow and should therefore be performed as infrequently and quickly as possible. Furthermore, the concurrent operation of the garbage collector may slow down the application threads

due to the amount of system resources it needs. Thus, an improved concurrent garbage collection technique that minimizes the frequency and duration of “stop-the-world” garbage collection while requiring less system resources would be advantageous.

5

SUMMARY OF THE INVENTION

In one aspect of the present invention a method is provided for collecting garbage in a computing environment, the method including a) tracing a root object to any of its reachable objects in a population of objects, b) marking any of the objects referred to in step a), c) unmarking a marked card including any of the objects, d) 10 tracing any marked object on the unmarked card to an unmarked referent object of the marked object, e) marking the unmarked referent object, f) tracing the referent object marked in step e) to any of its reachable objects, g) marking any of the objects referred to in step f), h) tracing any unmarked root object referent to any of its reachable objects, i) marking any of the objects referred to in step h), j) performing any of steps c) – g), 15 and k) designating any unmarked object in the population of objects as available for reallocation, where any of steps a) – g) are performed upon the population of objects concurrently with the operation of a mutator upon the population of objects within the computing environment, and where any of steps h) – k) are performed upon the population of objects while no mutator operates upon the population of objects within 20 the computing environment.

In another aspect of the present invention the method further includes marking the card if the mutator modifies an object pointer of an object in the card.

In another aspect of the present invention any of steps a) – g) are performed concurrently.

25

In another aspect of the present invention any of steps h) – j) are performed concurrently.

In another aspect of the present invention either of steps a) and f) are performed for a given object only if the card to which the object belongs is not marked.

30

In another aspect of the present invention the method further includes marking the card only if there is at least one marked object already on the card.

In another aspect of the present invention the method further includes periodically unmarking any marked card that does not contain at least one of the marked objects.

5 In another aspect of the present invention the method further includes designating any of the objects as “new”, and deferring the tracing of the “new” objects during any cycle of a plurality of cycles during which any of steps a) – g) are performed.

10 In another aspect of the present invention the designating as “new” step is performed if the object is part of an allocation cache from which objects are currently being allocated.

In another aspect of the present invention the method further includes periodically unmarking any marked card containing only “new” objects, and removing the “new” objects’ “new” designation.

15 In another aspect of the present invention the periodically unmarking and removing steps are performed if the object is part of an allocation cache from which objects are not currently being allocated.

20 In another aspect of the present invention a method is provided for collecting garbage in a computing environment, the method including a) tracing a root object to any of its reachable objects in a population of objects, b) marking any of the objects referred to in step a), c) unmarking a marked card including any of the objects, d) tracing any marked object on the unmarked card to an unmarked referent object of the marked object, e) marking the unmarked referent object, f) tracing the referent object marked in step e) to any of its reachable objects, g) marking any of the objects referred to in step f), h) tracing any unmarked root object referent to any of its reachable
25 objects, i) marking any of the objects referred to in step h), j) performing any of steps c) – g), and k) designating any unmarked object in the population of objects as available for reallocation, where either of steps a) and f) are performed for a given object only if the card to which the object belongs is not marked, where any of steps a) – g) are performed upon the population of objects concurrently with the operation of a mutator
30 upon the population of objects within the computing environment, and where any of steps h) – k) are performed upon the population of objects while no mutator operates upon the population of objects within the computing environment.

In another aspect of the present invention the method further includes marking the card if the mutator modifies an object pointer of an object in the card.

In another aspect of the present invention any of steps a) – g) are performed concurrently.

5 In another aspect of the present invention any of steps h) – j) are performed concurrently.

In another aspect of the present invention a method is provided for collecting garbage in a computing environment, the method including a) tracing a root
10 object to any of its reachable objects in a population of objects, b) marking any of the objects referred to in step a), c) unmarking a marked card including any of the objects, d) tracing any marked object on the unmarked card to an unmarked referent object of the marked object, e) marking the unmarked referent object, f) tracing the referent object marked in step e) to any of its reachable objects, g) marking any of the objects referred to in step f), h) tracing any unmarked root object referent to any of its reachable
15 objects, i) marking any of the objects referred to in step h), j) performing any of steps c) – g), and k) designating any unmarked object in the population of objects as available for reallocation, where prior to the unmarking step c) the card is marked only if there is at least one marked object already on the card, where any of steps a) – g) are performed upon the population of objects concurrently with the operation of a mutator upon the
20 population of objects within the computing environment, and where any of steps h) – k) are performed upon the population of objects while no mutator operates upon the population of objects within the computing environment.

In another aspect of the present invention the method further includes marking the card if the mutator modifies an object pointer of an object in the card.

25 In another aspect of the present invention any of steps a) – g) are performed concurrently.

In another aspect of the present invention any of steps h) – j) are performed concurrently.

In another aspect of the present invention a method is provided for
30 collecting garbage in a computing environment, the method including a) tracing a root object to any of its reachable objects in a population of objects, b) marking any of the objects referred to in step a), c) unmarking a marked card including any of the objects,

d) tracing any marked object on the unmarked card to an unmarked referent object of the marked object, e) marking the unmarked referent object, f) tracing the referent object marked in step e) to any of its reachable objects, g) marking any of the objects referred to in step f), h) tracing any unmarked root object referent to any of its reachable objects, i) marking any of the objects referred to in step h), j) performing any of steps c) – g), k) designating any unmarked object in the population of objects as available for reallocation, and l) prior to performing any of steps a) – g), periodically unmarking any marked card that does not contain at least one of the marked objects, where any of steps a) – g) are performed upon the population of objects concurrently with the operation of a mutator upon the population of objects within the computing environment, and where any of steps h) – k) are performed upon the population of objects while no mutator operates upon the population of objects within the computing environment.

In another aspect of the present invention the method further includes marking the card if the mutator modifies an object pointer of an object in the card.

In another aspect of the present invention any of steps a) – g) are performed concurrently.

In another aspect of the present invention any of steps h) – j) are performed concurrently.

In another aspect of the present invention a method is provided for collecting garbage in a computing environment, the method including a) tracing a root object to any of its reachable objects in a population of objects, b) marking any of the objects referred to in step a), c) unmarking a marked card including any of the objects, d) tracing any marked object on the unmarked card to an unmarked referent object of the marked object, e) marking the unmarked referent object, f) tracing the referent object marked in step e) to any of its reachable objects, g) marking any of the objects referred to in step f), h) tracing any unmarked root object referent to any of its reachable objects, i) marking any of the objects referred to in step h), j) performing any of steps c) – g), k) designating any unmarked object in the population of objects as available for reallocation, and l) during any cycle of a plurality of cycles during which steps a) – g) are performed designating any of the objects as “new”, and deferring the tracing of the “new” objects, where any of steps a) – g) are performed upon the population of objects concurrently with the operation of a mutator upon the population of objects within the

computing environment, and where any of steps h) – k) are performed upon the population of objects while no mutator operates upon the population of objects within the computing environment.

5 In another aspect of the present invention the method further includes marking the card if the mutator modifies an object pointer of an object in the card.

In another aspect of the present invention any of steps a) – g) are performed concurrently.

In another aspect of the present invention any of steps h) – j) are performed concurrently.

10 In another aspect of the present invention the designating as “new” step is performed if the object is part of an allocation cache from which objects are currently being allocated.

In another aspect of the present invention the method further includes periodically unmarking any marked card containing only “new” objects, and removing
15 the “new” objects’ “new” designation.

In another aspect of the present invention the periodically unmarking and removing steps are performed if the object is part of an allocation cache from which objects are not currently being allocated.

In another aspect of the present invention a system is provided for collecting
20 garbage in a computing environment, the system including a) means for tracing a root object to any of its reachable objects in a population of objects, b) means for marking any of the objects referred to in a), c) means for unmarking a marked card including any of the objects, d) means for tracing any marked object on the unmarked card to an unmarked referent object of the marked object, e) means for marking the unmarked
25 referent object, f) means for tracing the marked referent object marked in e) to any of its reachable objects, g) means for marking any of the objects referred to in f), h) means for tracing any unmarked root object referent to any of its reachable objects, i) means for marking any of the objects referred to in h), and j) means for designating any unmarked object in the population of objects as available for reallocation, where any of means a) –
30 g) operate upon the population of objects concurrently with the operation of a mutator upon the population of objects within the computing environment, and where any of

means h) – j) operate upon the population of objects while no mutator operates upon the population of objects within the computing environment.

In another aspect of the present invention the system further includes means for marking the card if the mutator modifies an object pointer of an object in the card.

5 In another aspect of the present invention any of means a) – g) operate concurrently.

In another aspect of the present invention any of means h) – i) operate concurrently.

10 In another aspect of the present invention either of tracing means a) and f) are operative to trace a given object only if the card to which the object belongs is not marked.

In another aspect of the present invention the system further includes means for marking the card only if there is at least one marked object already on the card.

15 In another aspect of the present invention the system further includes means for periodically unmarking any marked card that does not contain at least one of the marked objects.

In another aspect of the present invention the system further includes means for designating any of the objects as “new”, and means for deferring the tracing of the “new” objects during any cycle of a plurality of cycles during which any of means a) –
20 g) operate.

In another aspect of the present invention the means for designating as “new” is operative if the object is part of an allocation cache from which objects are currently being allocated.

25 In another aspect of the present invention the system further includes means for periodically unmarking any marked card containing only “new” objects, and means for removing the “new” objects’ “new” designation.

In another aspect of the present invention the means for periodically unmarking and the means for removing are operative if the object is part of an allocation cache from which objects are not currently being allocated.

30 In another aspect of the present invention a system is provided for collecting garbage in a computing environment, the system including a) means for tracing a root object to any of its reachable objects in a population of objects, b) means for marking

any of the objects referred to in a), c) means for unmarking a marked card including any of the objects, d) tracing any marked object on the unmarked card to an unmarked referent object of the marked object, e) marking the unmarked referent object, f) means for tracing the marked referent object in e) to any of its reachable objects, g) means for marking any of the objects referred to in f), h) means for tracing any unmarked root object referent to any of its reachable objects, i) means for marking any of the objects referred to in h), and j) means for designating any unmarked object in the population of objects as available for reallocation, where either of tracing means a) and f) trace a given object only if the card to which the object belongs is not marked, where any of means a) – g) operate upon the population of objects concurrently with the operation of a mutator upon the population of objects within the computing environment, and where any of means h) – j) operate upon the population of objects while no mutator operates upon the population of objects within the computing environment.

In another aspect of the present invention a system is provided for collecting garbage in a computing environment, the system including a garbage collector including a) means for tracing a root object to any of its reachable objects in a population of objects, b) means for marking any of the objects referred to in a), c) means for unmarking a marked card including any of the objects, d) tracing any marked object on the unmarked card to an unmarked referent object of the marked object, e) marking the unmarked referent object, f) means for tracing the marked referent object in e) to any of its reachable objects, g) means for marking any of the objects referred to in f), h) means for tracing any unmarked root object referent to any of its reachable objects, i) means for marking any of the objects referred to in h), and j) means for designating any unmarked object in the population of objects as available for reallocation, and a mutator operative to mark the card only if there is at least one marked object already on the card, where any of means a) – g) operate upon the population of objects concurrently with the operation of a mutator upon the population of objects within the computing environment, and where any of means h) – j) operate upon the population of objects while no mutator operates upon the population of objects within the computing environment.

In another aspect of the present invention a system is provided for collecting garbage in a computing environment, the system including a) means for tracing a root

object to any of its reachable objects in a population of objects, b) means for marking any of the objects referred to in a), c) means for unmarking a marked card including any of the objects, d) tracing any marked object on the unmarked card to an unmarked referent object of the marked object, e) marking the unmarked referent object, f) means
5 for tracing the marked referent object in e) to any of its reachable objects, g) means for marking any of the objects referred to in f), h) means for tracing any unmarked root object referent to any of its reachable objects, i) means for marking any of the objects referred to in h), j) means for designating any unmarked object in the population of objects as available for reallocation, and k) means for periodically unmarking any
10 marked card that does not contain at least one of the marked objects, where any of means a) – g) operate upon the population of objects concurrently with the operation of a mutator upon the population of objects within the computing environment, and where any of means h) – j) operate upon the population of objects while no mutator operates upon the population of objects within the computing environment.

15 In another aspect of the present invention a system is provided for collecting garbage in a computing environment, the system including a) means for tracing a root object to any of its reachable objects in a population of objects, b) means for marking any of the objects referred to in a), c) means for unmarking a marked card including any of the objects, d) tracing any marked object on the unmarked card to an unmarked
20 referent object of the marked object, e) marking the unmarked referent object, f) means for tracing the marked referent object in e) to any of its reachable objects, g) means for marking any of the objects referred to in f), h) means for tracing any unmarked root object referent to any of its reachable objects, i) means for marking any of the objects referred to in h), j) means for designating any unmarked object in the population of
25 objects as available for reallocation, k) means for designating any of the objects as “new”, and l) means for deferring the tracing of the “new” objects, where any of means a) – g) operate upon the population of objects concurrently with the operation of a mutator upon the population of objects within the computing environment, and where any of means h) – j) operate upon the population of objects while no mutator operates
30 upon the population of objects within the computing environment.

In another aspect of the present invention a computer program is provided, embodied on a computer-readable medium, the computer program including a) a first

code segment operative to trace a root object to any of its reachable objects in a population of objects, b) a second code segment operative to mark any of the objects referred to in a), c) a third code segment operative to unmark a marked card comprise any of the objects, d) a fourth code segment operative to trace any marked object on the
5 unmarked card to an unmarked referent object of the marked object, e) a fifth code segment operative to mark the unmarked referent object, f) a sixth code segment operative to trace the marked referent object in e) to any of its reachable objects, g) a seventh code segment operative to mark any of the objects referred to in f), h) a eighth code segment operative to trace any unmarked root object referent to any of its
10 reachable objects, i) a ninth code segment operative to mark any of the objects referred to in h), and j) a tenth code segment operative to designate any unmarked object in the population of objects as available for reallocation, where any of code segments a) – g) operate upon the population of objects concurrently with the operation of a mutator upon the population of objects within the computing environment, and where any of
15 code segments h) – k) operate upon the population of objects while no mutator operates upon the population of objects within the computing environment.

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention will be understood and appreciated more fully from
20 the following detailed description taken in conjunction with the appended drawings in which Figs. 1 - 7 are simplified flowchart illustrations of concurrent garbage collection methods.

DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

Reference is now made to Fig. 1, which is a simplified flowchart illustration
25 of a concurrent garbage collection method. In the method of Fig. 1 a garbage collector (GC) is run in parallel with one or more mutators. During concurrent operation, as the mutator modifies an object pointer it marks the card where the object is found. Marking, as described herein, may be performed using any known means, such as by
30 setting an object bit or byte value designated for GC marking purposes, or by storing objects' marks in a table. Preferably, no objects or cards are in a marked state prior to operation of the GC and the mutators. If the object spans cards, the mutator may mark

the card where the object begins or, alternatively, where the modified pointer is located. Meanwhile, the GC forms an object connectivity graph by tracing root objects to their referents and their descendents, marking all reachable objects.

Once the object connectivity graph has been traced, or while it is being
5 traced, card cleaning may be performed as follows. For each marked card, the GC unmarks the card and traces each marked object on the card to the object's direct referents only, marking each unmarked referent. Once all cards have been cleaned, or during the card cleaning process, the referent objects marked during card cleaning are traced to all their reachable objects, and all such objects are marked. Card cleaning and
10 object connectivity graph tracing may be performed one or more times for a predetermined number of iterations or until free heap memory is exhausted.

Concurrent operation of the GC and the mutators ends with the mutator processes being suspended. The GC then traces all unmarked root object referents to their referents and their descendents, marking all reachable objects and performs a card
15 cleaning pass. All unmarked objects may then be reclaimed, such as by listing them on a list of free objects available for reallocation. Concurrent operation of the GC and the mutators may then begin again, preferably once all objects and card marks are removed or reset.

Reference is now made to Fig. 2, which is a simplified flowchart illustration
20 of a concurrent garbage collection method. The method of Fig. 2 is similar to the method of Fig. 1 with the notable exception, shown in bold lettering, that during concurrent operation of the GC and the mutator, object tracing (where indicated by an asterisk) is performed for a given object only if the card to which the object belongs is not marked. This eliminates the need for the GC to trace marked objects on marked
25 cards both during the object tracing from the root objects, as well as during card cleaning. This also delays the GC's tracing from the most active cards (i.e., those marked most often by the mutator), which may reduce the number of cache misses due to concurrent access by the GC and the mutator to the same memory locations, as well as reduce the amount of the "floating garbage," unused memory which can not be
30 reclaimed by the GC.

Reference is now made to Fig. 3, which is a simplified flowchart illustration of a concurrent garbage collection method. The method of Fig. 3 is similar to the

method of Fig. 1 with the notable exception, shown in bold lettering, that the mutator only marks a card if there is at least one marked object already on the card. This reduces the amount of marked cards to be cleaned by the GC, thereby reducing the GC's work, shortening the duration of the stop-the-world phase, and reducing the number of cache misses during the concurrent operation of the GC and the mutators. The method of Fig. 3 may be combined with the method of Fig. 2, as is shown in Fig. 4.

Reference is now made to Fig. 5, which is a simplified flowchart illustration of a concurrent garbage collection method. The method of Fig. 5 is similar to the method of Fig. 1 with the notable exception, shown in bold lettering, that according to a predetermined schedule, and at any time relative to the other functions performed by the GC, the GC periodically unmarks marked cards that contain no marked objects. This eliminates the need for the mutator to check whether a card contains marked objects before the mutator marks the card. The method of Fig. 5 may be combined with the method of Fig. 2, as is shown in Fig. 6.

Reference is now made to Fig. 7, which is a simplified flowchart illustration of a concurrent garbage collection method. The method of Fig. 7 is similar to the method of Fig. 1 with several notable exceptions, shown in bold lettering. In the method of Fig. 7, objects that have recently been created by the mutator may be designated as "new" objects. For example, in the JAVA environment a mutator may be allocated a portion of memory known as an "allocation cache" that spans several cards from which the mutator allocates memory for objects it creates. Once the mutator has used up all of the allocation cache, it may receive another allocation cache, and so on. Thus, an object may be designated as "new" while it is part of an allocation cache from which the mutator is currently allocating memory for objects. Conversely, an object is no longer "new" if it is not part of an allocation cache from which the mutator is currently allocating memory for objects. The mutator may designate objects as "new" using any known marking technique.

Periodically, such as each time the mutator receives a new allocation cache, the mutator unmarks marked cards containing only "new" objects and removes the objects' "new" designation. For its part the GC defers the tracing of "new" objects, and traces them only once they lose their "new" designation. This reduces the number of marked cards to be cleaned by the GC, and, by delaying the tracing of newly allocated

objects, reduces the number of cache misses and floating garbage. The method of Fig. 7 may be combined with any of the methods described hereinabove.

It is appreciated that one or more of the steps of any of the methods described herein may be omitted or carried out in a different order than that shown, without departing from the true spirit and scope of the invention.

While the methods and apparatus disclosed herein may or may not have been described with reference to specific computer hardware or software, it is appreciated that the methods and apparatus described herein may be readily implemented in computer hardware or software using conventional techniques.

While the present invention has been described with reference to one or more specific embodiments, the description is intended to be illustrative of the invention as a whole and is not to be construed as limiting the invention to the embodiments shown. It is appreciated that various modifications may occur to those skilled in the art that, while not specifically shown herein, are nevertheless within the true spirit and scope of the invention.